5-6-2024

Artificial Intelligence and Music: Analysis of Music Generation Techniques Via Deep Learning and the Implications of AI in the Music Industry

David Bryce

Bryant University

HONORS THESIS

Artificial Intelligence and Music: Analysis of Music Generation Techniques via Deep Learning and the Implications of AI in the Music Industry

By David Bryce

ADVISOR • Tingting Zhao

EDITORIAL REVIEWER • Joan Zaretti

Submitted in partial fulfillment of the requirements for graduation
with honors in the Bryant University Honors Program
April 2024

Table of Contents

Artificial Intelligence and Music: Analysis of Music Generation Techniques via Deep Learning and the Implications of AI in the Music Industry
Honors Thesis for David Bryce

ABSTRACT

The use of artificial intelligence (AI) is quickly gaining relevancy in creative fields, and its emergence into the music industry comes with many unique implications. This paper examines the technical processes of creating music with AI and machine learning, the relationship between music and emotion, and finally the implications and ethical considerations for AI generated music in creative industries. As part of this project, a generative deep learning model (Music Variational Autoencoder) is explored and applied to generate music using a pre-trained training set of piano rolls. The AI reconstructions are based on self-made 4 measure electronic instrumental tracks. 46 machine learning students then take a survey to blindly compare the human generated tracks with the AI generated tracks to see if they can tell the difference. There are two trials of this survey, with a presentation on MusicVAE given in between the two trials. Exploratory analysis indicates that music experience is correlated with increased ability to distinguish AI generated music. Chi-Square Tests are then conducted for each set in each trial with a null hypothesis stating that the chance of guessing if a song is AI generated is equal to 50%. These results indicate that the null hypothesis cannot be rejected for the first trial, but that it is rejected for the second trial with the addition of the presentation in between.

INTRODUCTION

In the last few decades, the music industry has radically changed due to the introduction of electronic techniques such as sampling, remixing, and other new forms of performance (Réveillac, 2019, p. 11). The definition of what it means to be a musician has fundamentally changed as performers and DJs of all kinds have gained popularity. One of the latest advancements in the music world involves the creation of music via artificial intelligence (AI). This suggests a future of non-human musical artists, with some human input that will most likely be needed. The topic I have been investigating is the rise of AI in the music industry and how it has changed the role musicians can play in society and whether it can ever serve the same functions as music created by humans. This is important to the field because advances such as AI and deep learning have significant impact on a variety of industries but are recently beginning to infiltrate creative ones.

LITERATURE REVIEW

Introduction

Artificial Intelligence has made a drastic impact on different domains including business, art and now music. Artificial Intelligence (AI) is defined as the capability of machines to perform cognitive functions associated with human intelligence such as reasoning, learning, and perceiving (Berente, 2019, p. 1). AI can be categorized into a scientific (theoretical) side and an engineering (application) side, but music generation via AI incorporates both aspects (Roads, 1985, p. 1). Melodies/songs can now be created with AI and it is still unknown how this will affect the need for musicians and if these melodies will be good enough to use in commercials, movies, or even performances. This is further complicated due to the idea that music is not only defined by just its melody and lyrics, but also by the emotion that the artist puts into it and the emotions that listeners feel. The literature presented in this paper explores the technical aspects of AI music generation as well as its implications, the complex relationship between music and emotion, and the ethical dilemmas of the use of AI in the music industry.

Machine Learning Techniques for Music Creation

Machine learning techniques are becoming increasingly relevant in the creation of music and are already emerging in the music industry and other industries that require music technology. This study focuses on how these techniques are being used to create music and the implications these methods will have. Within this topic, some of the literature focuses on the specific and intricate machine learning methods used to create music and how machine learning can be utilized in a creative manner. Fiebrink (2018) defines machine learning as "the capacity of a computational system to learn structures from datasets in order to make predictions on newly seen data" (p. 1). Fiebrink explains that this allows musicians to teach computers to learn a specific style or even "break the rules" and explore the limitations of what is possible.

There are two main ways in which machine learning can occur: supervised learning and unsupervised learning. In a machine learning model, there is normally a training and testing set. In supervised machine learning, the training set features pairs of inputs and outputs for the model to associate together. "For example, consider a musician who would like to associate different hand positions, captured by a video camera, to different notes played by a computer. The musician can construct a training set by recording several examples of a first-hand position and labelling each one with the desired note, for instance "A#". She can then record examples of a second hand position, labelling each with another note, for instance "F". The training process learns what distinguishes an "A#" hand position from an "F" hand position. After training is complete, the musician can perform these two hand positions and use the trained model to label them as "A#" or "F" (Fiebrink, 2018, p. 4). Under an unsupervised learning context, however, only input data is given. There is no output/response variable for the model to predict but rather, the task is to learn the latent structure about the feature input. In a musical context, one may use unsupervised learning try to create a musical sample similar to the inputted samples when there is no sample output to guide the result. This relates to stylistic music generation, which is a method of generating melody that requires human control over features such as rhythm, genre, pitch, and structure. Jiang (2019) explains how researchers have already started to develop models for music generation, such as MusicVAE and MidiNet (Roberts, 2018; Yang 2017). These both aim to create new music without human guidance. However, Jiang also states that humans must stay in charge of some stylistic properties to develop a successful generative model that can answer questions such as how to create a similar melody with a specific sample. In a study using three different methods of generating computerized music, Jiang finds that even the most successful method (Long Short-term Memory Networks) has major issues such as less model accuracy when dealing with longer music. The process of unraveling the structure and content of the song is unsupervised, so effectiveness is hard to guarantee with this lack of supervision.

<u>Implications of AI in the Music Industry</u>

The literature also explores the future implications of AI generated music in the music industry and other industries. The full capabilities of AI generated music are still under development and not yet known, and there are many different opinions on the extent it will be useful in society. Frid (2021) suggests that there are specific ways to categorize the many different types of AI generated musical creations. "When it comes to computer-based music generation based on deep learning, two main types of systems can be distinguished: 1) autonomous music-making systems aimed at creation of original music for commercials and documentaries (e.g. Amper Music [52] and Jukedeck [36]), and 2) computer based environments based on assisting human musicians (e.g. FlowComposer [59] and OpenMusic [9]) [14]. To our knowledge, only one software solution using AI to generate music for video exists: Muzeek" (Frid, 2020, p.3). Bonnici (2021) states that "Music must be performed expressively to be engaging". This means that generating a melody alone is not enough. For AI generated music to be captivating, human expression must be simulated or at least incorporated. However, Bonnici mentions that there are other industries beyond the music industry for which AI generated music may be a perfect fit. For example, AI generated Music is gaining relevancy in health care, as this type of music can be created for very specific purposes with therapeutic constraints in mind. "On the use of AI for Generation of Functional Music to Improve Mental Health (Williams et al.) uses machine learning to create music targeting a specific physiological response. This work suggests a new direction for the evaluation of music generation systems as well as future applications such as games and health care" (Bonnici, 2021, p. 2). This suggests that computer generated music may be best suited for less creative musical tasks in specific industries such as healthcare that require specific features. Dredge (2017) argues that AI generated music does not have to be better or even as good as human generated music to fit a specific purpose, especially for people who do not have large budgets. "Christopher Nolan is not going to stop working with Hans Zimmer any time soon," says Cliff Fluet, partner at London law firm Lewis Silkin, who works with several AI music startups. However, he believes that AI generated music could be very useful for youtubers who are making short films or those trying to avoid copyright violations (Dredge, 2017, p. 2). The question remains whether AI generated music ever has the potential to replace the need for human musicians. Music Industry consultant Mark Mulligan believes that AI might never be able to make music that will move people like human music does. ""Because making music that moves people -- to jump up and dance, to cry, to smile -- requires triggering emotions and it takes an understanding of emotions to trigger them," he says. "If AI can learn to at least mimic human emotions then that final frontier may be breached. But that is a long, long way off."" (Dredge, 2017, p.3).

<u>Music and Emotion</u>
The complex relationship between music and emotion is also featured in the literature. Understanding this relationship will help to determine if AI can ever truly simulate human emotion in a way that creates moving music. David Carr (2004) does not describe one simple relationship between music and emotion. He describes multiple complex theories about a potential relationship but ends by arguing that music helps us connect to our own emotions more so than just inherently representing a specific emotion. Carr states that "by listening to "Carmen" or "Othello" audiences may come to learn more about their own feelings of anguish or jealousy". This promotes the idea that music helps us learn about our emotions, and that there is a certain complexity between music and emotion that makes it hard to define. Holt (2007) argues that music can be intertwined with strong messages that produce a meaningful and emotional response. For example, many popular songs incorporate ideas surrounding racism, sexism, and politics and gain power in the struggles against these social problems. This suggests that emotion and meaning is derived from the powerful and complex social and

cultural concepts that specific songs represent. Kawakami (2013) defines two distinct types of emotion in regard to music: perceived emotion and felt emotion. Perceived emotion is the emotion intended by the musical artist and felt emotion is the emotion that the listener feels. The literature describes an experiment using 12 males and 12 females with similar levels of music experience to measure the correspondence between perceived and felt emotions. Two novel musical stimuli were created so that no participants would recognize a song, and they were asked to compare the two with a series of questions. Kawakami concludes by explaining that "we hypothesized that perceived emotion would not necessarily correspond to felt emotion, and we investigated which musical structures contributed to this difference. Participants with high levels of music experience who listened to dissonant music and music in a minor key judged the perceived emotion to be more unpleasant, but the emotions they experienced were not as unpleasant as the ones they perceived. This finding may contribute to the study about the fascination of sad music". This adds further complexity to the relationship between music and emotion as it suggests that there can be a fascination around negative emotions in music that can lead to listeners feeling more connected with a track.

Ethical Issues of AI Music

The last major topic in this literature revolves around the ethical issues surrounding AI generated music and its implications. A large ethical issue with this type of music involves copyright. There is a debate on if AI generated music warrants copyright protection at all. "According to the Court of Justice of the European Union (CJEU), a work is considered original when it is the expression of the author's own intellectual creation and his/her free creative choices, the author's personality, or the author's personal touch. Considering this, several scholars conclude that under present law, autonomously AI-generated works might not be eligible for copyright protection" (Sturm, 2019, p.4). Sturm argues that humans will still have a crucial involvement in the generation of AI generated music, as the AI system is essentially helping the artist achieve their vision. In order for a project to be original (which is important for copyright purposes), there needs to be a substantial creative contribution by a human which can still be done with AI generated works. This may not apply if in the future there becomes a way for AI to autonomously generate songs, but until then copyright protection may still apply. In addition to copyright, there are other ethical issues involving AI and music. Morreale (2021) describes the issue of musician redundancy regarding commercial music platforms that use an AI based recommendation system (such as Spotify, Apple Music, etc.). He defines this as the "long tail problem" (p.107). This represents the idea that a small subset of musicians are very popular while the majority are not given as much exposure. In turn, bias within AI methods exists in the music industry even without considering the generation of music with these methods. In regard to music generation, Artificial Intelligence Music (AIM), companies tend to avoid providing details about the training data sets used in their machine learning models. Data is also likely to include specific information about users and their behaviors and environment. Every time a user interacts with a specific song, their preferences and traits are being fed into a specific algorithm (p.108). Therefore, the ethical implications of AI methods in the music industry span much further than just music creation itself.

Literature Review Summarized

This literature demonstrates many techniques for generating music and melodies using AI and machine learning and addresses the limitations of current methods. It also gives a better understanding of potential relationships between music and emotion. Its strengths are that it gives details into specific methods as well as specific studies done in this field. However, its weaknesses are that it is mostly editorials and could use some more research beyond specific machine learning models. What is missing is an understanding of if AI will ever be able to produce music that can provoke the same reactions as humans. Real world examples of AI and music in action are also missing, as most of the studies featured in the literature are done on a smaller scale and are written in an overly technical manner. Finding research that could connect the research on machine learning to the research on emotion could help tie the literature together and make it easier to explore the potential of AI generated music. There is uncertainty of the potential of AI regarding music, and most of the literature does not make a strong claim around this issue. The next step is to apply an existing machine learning model or build a novel model if possible, to generate music and gain hands-on experience to understand the potential and limitations of the generative deep learning models on music.

Survey Results will be used to evaluate and compare the quality of the AI generated music with human generated music and suggest future directions of AI-generated music based on our study and survey results.

RESEARCH QUESTIONS

1. How is the rise of Artificial Intelligence changing the role that musicians can play in society and the way that people perceive music?
2. How are deep learning models used to generate music? What model structure has been used to allow for the capacity of deep learning in the music domain?
3. What are the roles that AI generated music will play in society in the future?
4. Will AI generated music ever be able to serve the same functions as human generated music?

METHODOLOGY/ETHICAL CONCERNS:

Track Creation

To generate AI Music for this project, the two necessary components are a deep learning model capable of creating AI regenerations of tracks, and the creation of tracks for the model to reconstruct. After exploring several state-of-the-art deep learning models in the literature to generate music, the unsupervised learning generative model of Music Variational Autoencoder is chosen. A pre-trained implementation of Music VAE from a bachelor's thesis in informatics at the Technical University of Munich will be used.

The music used as input into this model is self-composed using Garage Band. After inputting the melody midi files and the drum midi files into the Music Vae Implementation, a regenerated sample is created of specified length.

```
C:\Users\student\Downloads\tss19-VAE-music-generation>python
musicVAE.py reconstruct 2 -ptv -n 10 --start_sequence-
Sampled/manually_created_sequences/miditest.midi -s Sampled/AIMusicTest

loaded Checkpoint pianoroll_train_2bars_
1stride_tempo_computed_transposed

using device cpu
```

Note: This example is created a 2 measure reconstruction based on the midi file "miditest.midi" and saving the before and after results to a folder called AIMusic Test. It is loading a 2-measure piano roll file as the training set as this is a 2 measure reconstruction.

reconstruct_after.midi	11/10/2023 8:26 PM	MIDI Sequence	1 KB
reconstruct_before.midi	11/10/2023 8:25 PM	MIDI Sequence	1 KB

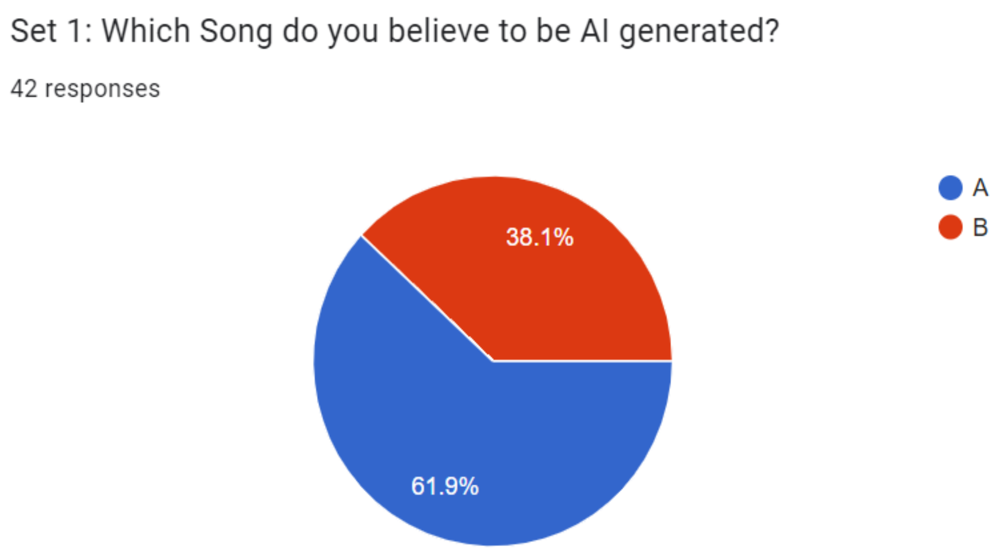
The reconstructed midi files are then put back into GarageBand and receive the same presets and effects as the original tracks did. This process is repeated 4 times so that there are 4 pairs of tracks (one human generated and one AI generated).

<u>Survey Experiment</u>

In the survey experiment, 3 Bryant University Machine Learning Classes (47 people total), both at an undergraduate and undergraduate level, listen to the 4 pairs of tracks (one human generated and one AI generated each) and attempt to distinguish which track is AI generated. Google Forms will be used for this survey. Each track is named with their set number and either A or B. For example, the first track is named "Track 1A". The AI generated track could be in either A or B for each set, and each set is independent from the last. Participants will also be asked their grade level (Undergraduate or Graduate) and their experience with music theory. The first survey attempt will be done with no prior information, but a presentation on Music VAE and strategies for identifying AI generated music is given in between the two trials. The second trial is then completed in the same format as the first and the data is automatically added to an excel spreadsheet upon submission.

Creation of Data Visualizations

Google Forms automatically creates pie charts for each question that shows the percentage of students that selected each answer.



This is helpful in immediately understanding which sets students were more easily able to pick out the AI generated track within. It is also helpful in comparing first trial results with second trial results.

Several bar graphs in Tableau are also created to demonstrate these three relationships:

1. Correct Answers and Experience with Music Theory
2. Correct Answers and Grade Level
3. Correct Answers and Trial Number

These visualizations will be shown and analyzed in the Results section.

Chi-Square Tests

To gage whether the results are significant, Chi-square tests must be conducted. This is used to determine whether two categorical variables are independent regarding the test statistic. Chi-Square tests were conducted for each song set with A and B being the categorical variables. For example, if the correct answer is A for Set 1 and 2 and 27 out of 46 selected A, then 27 would be the basis for the significance test given that there are 46 total records. The Chi Squared statistic is calculated, and then the p-value is calculated. If the p value is below

0.05, then the ability to guess the correct answer (A or B) is significant. Below are the formulas used for these calculations:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

- O_i represents the observed frequency for category i,
- E_i represents the expected frequency for category i, assuming the null hypothesis is true,
- The summation $\sum$ is over all categories.

$$\text{p-value} = 1 - F(\chi^2; df)$$

Df= degrees of Freedom

$F(\chi^2; df)$ = Cumulative Distribution Function

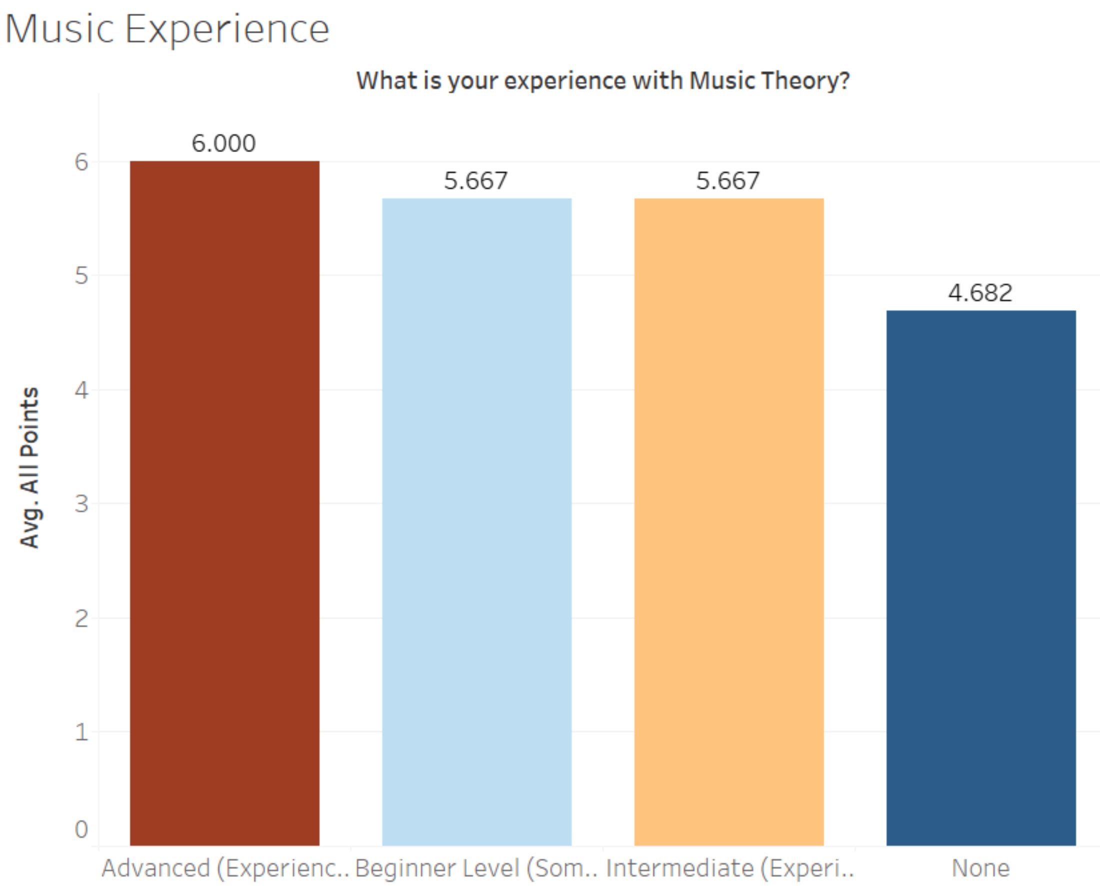
Cumulative Distribution Function (CDF) = probability that a random chi squared variable is <= the observed X^2

RESULTS AND DATA ANALYSIS:

After compiling the data from the 46 survey participants, the total number of guesses (# of participants * number of total guesses per person for all trials) is 368. Each correct answer is tallied as a point in the following visualizations. Therefore, the highest number of points anyone could receive is 8, only if they get every answer correct.

Data Visualizations

Figure 1: Music Experience vs Correct Guesses



This bar graph shows the average amount of points received for each category of music knowledge (Advanced, Beginner, Intermediate, None). This question was intended to help identify the relationship between music knowledge and experience and the ability to distinguish AI generated music. Those with advanced music knowledge (3 participants) on average made 6 correct guesses out of 8 questions (75% accuracy). Those with both beginner knowledge (12 participants) and intermediate knowledge (9 participants) both averaged 5.667 correct answers (70.84% accuracy). Lastly, those with no music knowledge (22 participants) only correctly guessed 4.682 out of 8 questions correctly on average (58.525% accuracy). Although sample size is not the same for each category, the data shows that those with advanced music experience are more likely to correctly identify AI generated music than those without musical experience. This suggests that certain technical musical elements distinguishable by musicians are either not present or noticeably different to those with a substantial music background.

Note: Here is the full distribution of participant Musical Experience

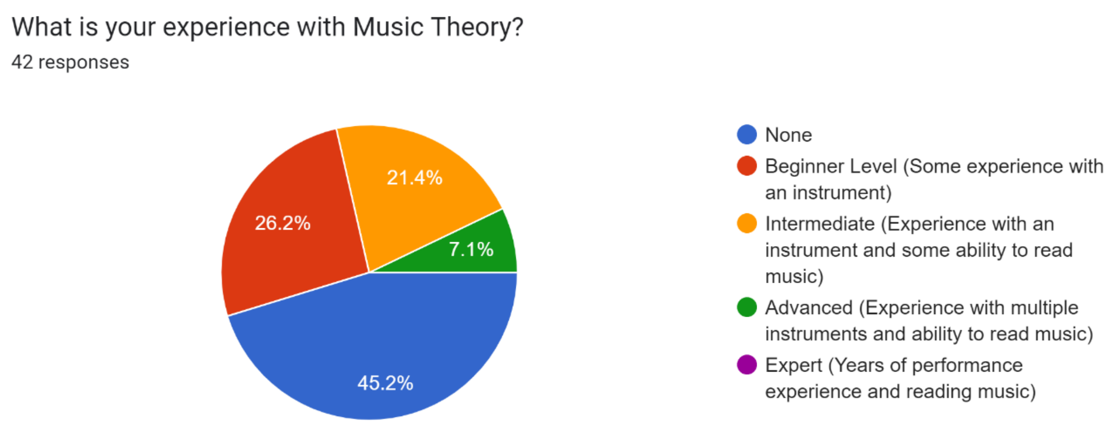
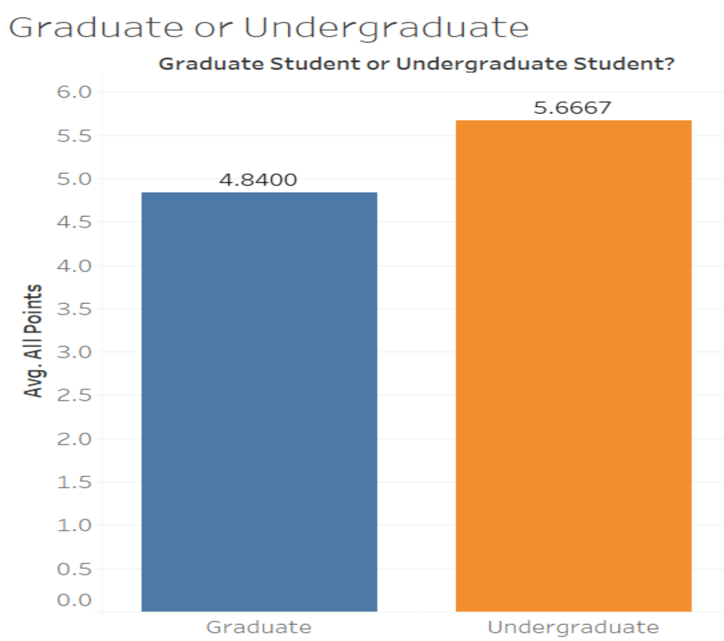
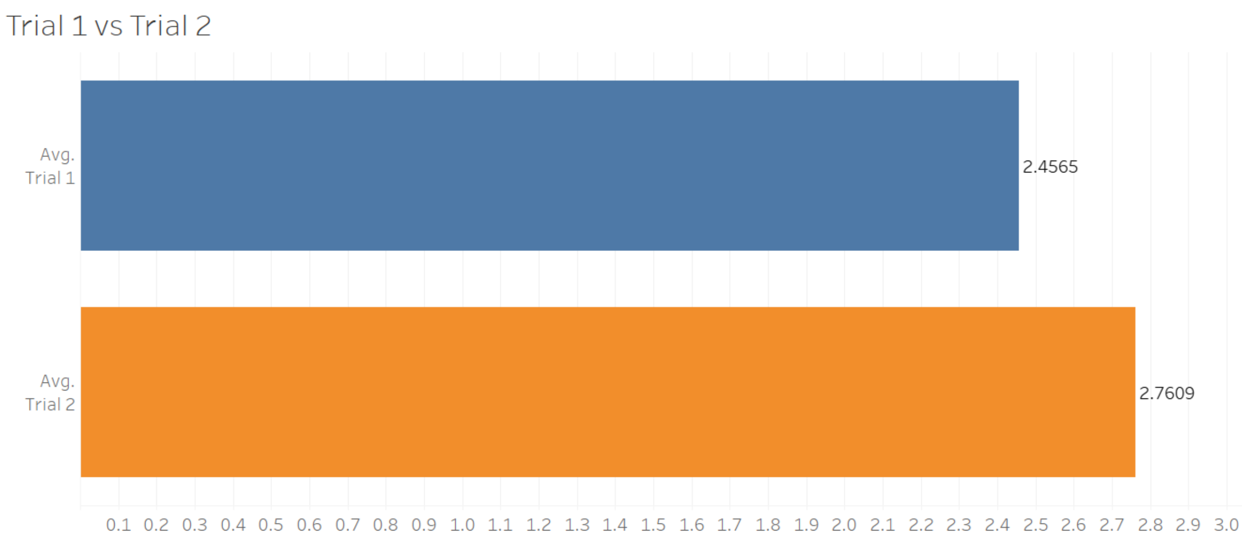


Figure 2: Education Level vs Correct Guesses



This bar graph shows the average amount of correct guesses categorized by Education Level (Graduate or Undergraduate). Out of the 46 participants, 25 were graduate level students and 21 were undergraduate level students. On average, undergraduate students were able to correctly guess the AI generated track 70.84% of the time (5.6667/8) while graduate students were only able to correctly identify the AI generated track 60.5% of the time. The data suggests that Bryant University undergraduate students taking Machine Learning are more likely to correctly identify AI generated music than those at the graduate level. Due to the small sample size, it is hard to generalize further than this.

Figure 3: Trial # vs Correct Guesses



This bar graph shows the average amount of correct guesses for each of the two trials. Trial 1 was given with no background information, while a presentation on MusicVAE and the process of music creation was given between Trial 1 and 2. Each trial had the same number of questions, and therefore the same number of responses (182 each). The highest point total for each trial is 4 points per person. On average, the participants performed better in Trial 2 (69.02% accuracy rate) than they did in Trial 1 (61.41% accuracy rate). This suggests that when educated about Music VAE and the process of generating AI music, students are more easily to identify AI generated tracks.

Chi-Square Tests Results

After compiling total correct answers (out of 46) for each question, the Chi Square tests were performed as outlined in the methodology. These tests show if the relationship between guessing correctly or incorrectly (A or B) for each question is due to random chance or if there are other factors at play. The null hypothesis is that there is a 50% chance of guessing the AI generated track correctly. If the p value is less than 0.05, we reject the null hypothesis and say that the results are not due to random chance.

- If a p value of less than 0.05 was obtained, that trial is highlighted in green.
- Set 4 Trial 2 is highlighted in red as it is a special case. This is the only trial where there are so few correct answers that the number is significant.

Figure 4: Significance Tests

	Set 1	Set 2	Set 3	Set 4
Trial 1 # correct	27	33	28	25
Trial 1 p value	0.238	0.0032	0.14	0.555
Trial 2 # correct	31	41	39	16
Trial 2 p value	0.018	0.00000011	0.0000024	0.029

For all the questions set in Trial 1, only 1 of the results out of 4 is significant. Therefore, only 25% of the question sets were marked as significant. For most of the results in trial 1, it cannot be proven that the outcomes are not due to random chance so the null hypothesis cannot be rejected.

However, in Trial 2, the results for 100% of the sets are significant. For sets 1-3, the number of correct answers increased significantly. In set 4, however, 9 less people were able to correctly identify the AI track than in Trial 1. The result for this set was not significant in Trial 1, but it became significant in Trial 2 due to the decrease in correct responses. This raises the question of why performance in Trial 2 improved for three of the sets but worsened for the last set. This will be explored further in the next section. The only changes between the first and second trials is that the participants had heard the songs once before, and they had received a presentation on Music Variational Autoencoder and musical differences between human and AI generated tracks. With these changes, the null hypothesis of the chance of guessing the AI song correctly being 50% is rejected for every set in the second trial. The addition of the presentation in between the trials is correlated with more significant results and a higher correct answer rate for 75% of the sets.

Musical Analysis

Set 4 is the only set where participants performed worse after receiving the Music VAE presentation. Within this presentation, the following helpful tips for distinguishing AI music were given.

1. **Notes and rhythms may sound out of place.**
2. **The track may feel scattered and have higher ranges of notes.**
3. **The human tracks may have more defined rhythmic patterns.**

The addition of this advice along with a basic understanding of MusicVAE correlated with higher correct guessing rate for every set except set 4. This indicates that the AI regeneration for set 4 did not follow the patterns outlined in the given tips. Due to the inherent randomness of the model under specified constraints, this is possible that Set 4 was able to more closely mimic human generated tracks. To further examine this, I transcribed both the human generated and AI generated tracks for Set 4

Figure 5: Transposed Set 4

Human Generated (Set 4):

AI Generated (Set 4):

(The AI generated track is only two measures as this is what was specified in the code for the track regeneration)

The human generated track is purposely designed to be simple and have repetitive, recognizable note patterns. It contains quarter notes and eighth notes, with simple but well thought out rhythms, and a harmonically pleasing melody. The AI generated track, while only two measures, starts with a repetitive 3 quarter note pattern, exactly the same as the human generated track starts. The entire first measure is the same between the tracks, except the 2nd eighth note in the AI generated track is lower than the previous note, while in the human generated track the 2nd eighth note is higher than the first eighth note. In the 2nd measure, the AI track deviates entirely from the human generated track with a slightly syncopated rhythm with a melody where the pitch is increasing steadily. The syncopated rhythm provides a more upbeat feel, while the changing melody also provides contrast from the repetitive first measure. This AI regeneration incorporates more technically complex rhythms and melodies than the human generated track, while still sounding just as pleasant and cohesive to the average ear. This same phenomenon did not occur for the other AI generated tracks, as can be seen below when comparing the sheet music for Set 3.

Figure 5: Transposed Set 3

Human Generated (Set 3):

AI Generated (Set 3):

For Set 3, 11 more students correctly identified the AI generated track in trial 2 than they did in trial 1. This set saw the highest improvement rate of any set between trials. Therefore, after receiving tips about distinguishing AI music and learning about MusicVAE, many more students were able to find the AI track for this set. Referring back to the tips explained at the beginning of this section, there is one note in the AI generated track that is clearly out of place both pitch-wise and rhythmically. This is the 16th note (A Natural) on the e of 2 (middle of the second measure). The previous two notes are a flat eighth notes separated by eighth rests, but the out of place note is only separated by a 16th note rest and is an A natural instead of an A flat. This note is unexpected as there is a strange pitch change and comes in slightly earlier than the previous two notes. This makes the syncopation feel sloppy and rushed. Therefore, as opposed to set 4, this follows the tips given in the presentation. It is also shows that with this model, AI tracks can have notes/rhythms out of place, but as we saw with Set 4, the AI generated patterns can occasionally be a subjective improvement over the human generated tracks.

CONCLUSIONS AND FUTURE RESEARCH

After conducting outside research, using MusicVAE to create AI generated music, and conducting an experiment to see if students could tell the difference, it is clear that AI music has usefulness in a variety of industries. MusicVAE functions by learning a latent space of notes and regenerating a track so that there are certain rhythm and pitch similarities. This can be very helpful to create melodies of a specific mood/genre to function as background music in commercials, tv shows, or events. However, the implementation of MusicVAE that was used for this project tended to create melodies with noticeable pitches and rhythms that sounded out of place. This phenomenon may not occur with models such as those used in Google's new Dream Track tool.

There were several limitations of this project, including the small sample size of 46. Every single person in the sample was a Bryant University student as well. It is hard to make a generalized claim with these results due to this sampling process. Additionally, this pre-trained model was implemented by one person, and advanced tuning options were not able to be explored. Certain phenomena noticed in these results will not be able to be generalized to include the results of other models.

One thing that remains clear is that human creativity and human musicians will still be useful despite present and future advancements in AI generated music. As we have noticed with the AI generated tracks in this project, the human ability to notice musical elements being out of place and tweak a track in an artistic manner is something that AI generated models cannot master. This is because the greatness of music is inherently subjective, although the trained human ear and even the untrained human ear to an extent, allows personal connections to be made with tracks in a way that a model can imitate but never replicate. There is no "correct way" for music to advance and progress in the future, and this is why the unpredictable creative minds of humans are irreplaceable. As stated previously, there are several industries where these tracks and melodies can function as background music, and the ability to create multiple similar tracks quickly with Music VAE is a distinct advantage. Ideas from AI generated melodies could also be a great starting point for artists to build there next big track from. However, it is very unlikely that there will be a world where AI generated songs become more popular or even as popular as human generated tracks. The fine details, the emotion, and the subjective beauty of music make it a distinctly human form of expression that AI methods cannot deeply infiltrate.

REFERENCES:

Berente, N., Gu, B., Recker, J., & Santhanam, R. (2019). Managing AI. MIS Quarterly. https://misq.umn.edu/skin/frontend/default/misq/pdf/CurrentCalls/ManagingAI.pdf

Bonnici, A., Dannenberg, R. B., Kemper, S., & Camilleri, K. P. (2021). Editorial: Music and AI. *Frontiers in Artificial Intelligence*, *4*, 651446. https://doi.org/10.3389/frai.2021.651446

Carr, D. (2004). Music, meaning, and emotion. *The Journal of Aesthetics and Art Criticism*, *62*(3), 225–234.

Dredge, S. (2017). AI and music: Will we be slaves to the algorithm? *The Observer*, 19.

Fiebrink, R. A., & Caramiaux, B. (2018). *The Machine Learning Algorithm as creative musical tool* (R. T. Dean & A. McLean, Eds.; Vol. 1). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190226992.013.23

Frid, E., Gomes, C., & Jin, Z. (2020). Music Creation by Example. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–13. https://doi.org/10.1145/3313831.3376514

Holt, F. (2007). *Genre in popular music*. University of Chicago Press. http://ebookcentral.proquest.com/lib/bryant/detail.action?docID=686254

Jiang, J. (2019). *Stylistic Melody Generation with Conditional Variational Auto-Encoder* [Class Project]. Carnegie Mellon University. https://www.cs.cmu.edu/~epxing/Class/10708-19/assets/project/final-reports/project8.pdf

Kawakami, A., Furukawa, K., Katahira, K., Kamiyama, K., & Okanoya, K. (2013). Relations

Between Musical Structures and Perceived and Felt Emotions. *Music Perception*, *30*(4), 407–417. https://doi.org/10.1525/mp.2013.30.4.407

Morreale, F. (2021). Where Does the Buck Stop? Ethical and Political Issues with AI in Music Creation. *Transactions of the International Society for Music Information Retrieval*, *4*(1), 105–113. https://doi.org/10.5334/tismir.86

Réveillac, J.-M. (2019). *Electronic music machines: The new musical instruments*. John Wiley & Sons, Incorporated. http://ebookcentral.proquest.com/lib/bryant/detail.action?docID=5761049

Roads, C. (1985). Research in music and artificial intelligence. *ACM Computing Surveys*, *17*(2), 163–190. https://doi.org/10.1145/4468.4469

Roberts, A., Engel, J., Raffel, C., Hawthorne, C., and Eck, D. (2018). A hierarchical latent vector model for learning long-term structure in music. https://arxiv.org/abs/1803.05428

Sturm, B. L. T., Iglesias, M., Ben-Tal, O., Miron, M., & Gómez, E. (2019). Artificial Intelligence and Music: Open Questions of Copyright Law and Engineering Praxis. *Arts*, *8*(3), Article 3. https://doi.org/10.3390/arts8030115

Yang, L.-C., Chou, S.-Y., and Yang, Y.-H. (2017). Midinet: A convolutional generative adversarial network for symbolic-domain music generation. https://arxiv.org/abs/1703.10847